

一筆描きされた絵からの感情認識

Emotion Recognition from One-Stroke Sketch

鈴木哲司
Satoshi Suzuki

橋浦康一郎
Koichiro Hashiura

能登谷淳一
Junichi Notoya

草苺良至
Yoshiyuki Kusakari

堂坂浩二
Kohji Dohsaka

秋田県立大学
Akita Prefectural University

This paper presents a method for recognizing emotion of the authors of one-stroke sketches. We collected 840 sketches from 35 authors in the experiment in which each author was requested to draw a one-stroke sketch representing an emotion. Emotions were grouped into eight classes based on the Plutchik's theory. We trained an emotion classifier from the collected data using the Support Vector Machine. We exploited a number of features extracted from the drawn image and the drawing process and we found that some features such as color (hue, saturation, and value) and HOG (Histogram of Oriented Gradients) were closely related to the author's emotion. The classification accuracy using 10-fold cross validation was 51.9%. When grouping emotions into two classes by assembling similar emotions into one group, the classification accuracy was 71.9%.

1. はじめに

人との会話を通して生活支援や学習支援などを行う会話ロボットや、人間同士の会話に介入し、会話を活性化させる会話ロボットの研究が盛んに行われている[1,2,3]。人間同士の会話においては、情報のやり取りだけでなく、感情のやり取りが重要な役割を果たす。人・ロボット間の会話においても、ロボットによる共感的な言語表現が人・ロボット間の多人数会話を活性化することが報告されている[3]。

そこで、本研究では、誰でも簡単に描くことができる一筆描きの絵という非言語コミュニケーション手段に着目し、一筆描きされた絵から描き手の感情を認識する手法を開発することを目的とする。人・ロボット間のやり取りにおいて、言葉だけでなく絵を使って感情を表現することができれば、親しみやすく、気軽に使える人・ロボットのコミュニケーション環境の実現が期待できる。例えば、言葉と絵を使った人間同士の会話の場にロボットが介入し、絵から認識した描き手の感情を用いて会話を盛り上げるロボットの開発が可能となるだろう。さらに、絵の描き手の感情認識は、人・ロボットのコミュニケーションにとどまらず、様々な情報システムに応用可能である。例えば、絵と文章を使って日々の出来事を記録する日記システムにおいて、ユーザの日々の感情変化を把握することや、感情カテゴリで日記を検索することが可能となる。

色と感情の間には密接な関係があることが知られており、山下の研究では色彩と人の感情との相関についての実験調査結果が報告されている[4]。実験では「喜・怒・哀・楽」の主観的な感情形容詞と、「寛」の意味で状態の判断を表す属性形容詞の計12形容詞が用いられた。結果として、喜・怒・哀・寛のそれぞれの感情からイメージされる色度はある共通した傾向を持つことが分かった。また、彩度・明度に関しても感情によって共通する点があることが分かっている。本研究では、この結果から、色相だけでなく彩度と明度についても考えることで感情の認識率の向上ができると期待し、実験データを取得する際に色は色相・彩度・明度の変更が行えるHSV色空間を使用する。

関連研究として、水を含んだ風景写真から感情を認識する

Dellagiacomaらの研究[5]がある。この研究では人が風景写真を見て知覚する感情を認識する認識器をサポートベクターマシンにより機械学習した。感情の分類はEkmanによる5分類(悲しみ、幸せ、嫌悪、恐れ、驚き)[6]を使用している。5分類の感情を用いて機械学習を行ったとき、分類正解率61.82%が得られている。また、ポジティブ(幸せ)とネガティブ(悲しみ・嫌悪・恐れ)の2分類にしたとき、分類正解率は80.56%であった。

この関連研究に対し、本研究はシンプルな一筆描きの絵を用いることに特徴がある。本研究では、人・ロボット間のコミュニケーションや絵を使った日記システムなどに絵の描き手の感情認識を適用することを意図しており、ここでは、一筆描きの絵のように、絵を使って容易に感情を表現できることが必要不可欠である。また、関連研究では写真を見た人が感じる感情を認識しているのに対し、本研究では絵の描き手の感情を認識する点が異なる。

2. 研究方法

2.1 手順

本研究の目的は一筆描きされた絵から描き手の感情を認識する認識器を開発することである。感情の分類にはPlutchikの分類を使用し、機械学習にはサポートベクターマシンを用いた。

研究の手順として、まず、描画データ収集実験によって大量の一筆描きの絵の描画データ(絵の画像データ、描画プロセスデータ)を人から収集する。次に、絵の描き手の感情を認識するのに有効な特徴量を選び出す。最後に、選んだ特徴量を用いて、感情認識器を描画データから学習する。

2.2 Plutchikによるの感情の分類

感情の分類としてPlutchikの分類[7]を用いた。心理学者であるPlutchikは感情には喜び、悲しみ、受容、嫌悪、恐れ、怒り、驚き、受容の8分類の基本感情があるとしている。また、それぞれの感情には対極の感情が存在し、類似した感情は隣接するという考えから図1に示すような感情の輪を提示している。その他にも感情の強度や、基本感情の両極同士を除いた2つの感情の組み合わせからなる混合感情についても述べられている。



図 1. Plutchik の感情の輪

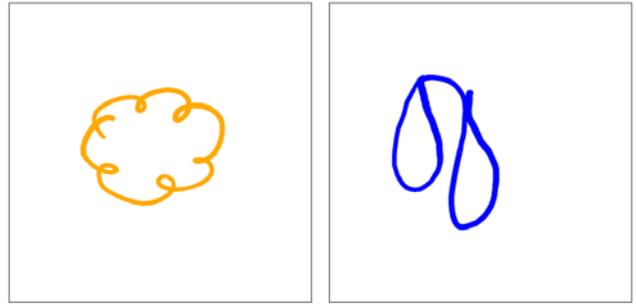


図 3. 一筆描きの感情の例(左:喜び, 右:悲しみ)

2.3 イラストツール

本研究では描画データ(絵の画像データ、描画プロセスデータ)を取得するためにイラストツール(図 2)を作成した。イラストツールの基本的な仕様を表 1 に示す。キャンバスサイズに対して表示するサイズを小さくしている理由は、絵を描く際にキャンバスの外に描かれてしまうと座標などの正確なデータが取得できないためである。イラストツールでは線の太さの違いによる感情の表現が行えるよう、筆圧の強弱によって筆の太さが変化するようにした。描いてもらう感情はイラストツールの右上に表示した。また、実験においては感情の強さを考えるため、感情の強さを表す大・中・小の文字を感情の右側に表示した。

色の選択はイラストツールの右下に表示したパレット(図 2-①)を使用する。一色を選択することで、選択した色の彩度と明度を変えることのできるパレット(図 2-②)が表示され、これから色を再び選択することで絵を描くことが可能となる。本研究では彩度値として 0%を用いていない。彩度値を 0%とすると、どのような色相や明度でも白-黒間の色となってしまうが、本研究では色として白-黒間の色を使用するので、これと区別するためである。明度値についても同様の理由で 0%を用いていない。明度値を 0%とすることで、どの色相や彩度でも黒色となるからである。イラストツールを用いて描かれた感情の絵の例を図 3 に示す。このような絵から感情を認識することを考える。

表 1. キャンバスの基本仕様

キャンバスサイズ	1000×1000[pixel]
キャンバス表示サイズ	800×800[pixel]
色の種類	19 色
明度の値	20%,40%,60%,80%,100%
彩度の値	20%,40%,60%,80%,100%

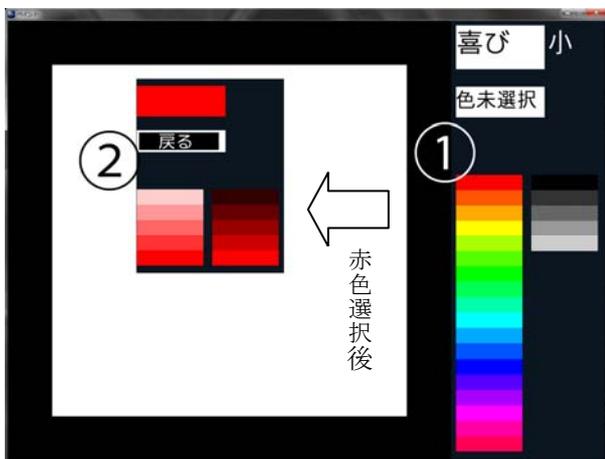


図 2. イラストツール

2.4 特徴量

機械学習に使用する学習データとして、本研究では一筆描きの絵の描画プロセスデータと絵の画像データから抽出した特徴量を使用する。

(1) 線を描く加速度の変化数

絵を描くときの加速度の増減が変化した回数の特徴量としたものである。例として、描く加速度が増加傾向から減少傾向に、減少傾向から増加傾向に切り替わるときにカウントとすることで求められる。

(2) 線を描く方向の変化数, X, Y 軸方向の割合

線を描く方向が下方向から上方向へ、右方向から左方向へというような変化をした回数の特徴量としたものであり、横(X 軸)方向と縦(Y 軸)方向を別々に求めている。また、X, Y 軸方向の割合とは X 軸方向と Y 軸方向の変化数の割合を示す。

(3) 筆圧の変化数

描いている時の筆圧が増加傾向から減少傾向に、減少傾向から増加傾向に切り替わったときにカウントすることで求められる特徴量である。

(4) 描き始め, 終わりの位置

キャンバスを 4×4 分割したときに絵を描き始めた, 終えた位置を $x=(1,2,3,4)$ と $y=(1,2,3,4)$ でそれぞれ表した特徴量である。

(5) 色(色相, 明度, 彩度)

絵を描くときに用いた色相, 明度, 彩度を表す特徴量である。色相は 0~340° を 20° 分割したものを色 1~色 18 にそれぞれに対応させた。また、黒色は色相 0°, 彩度 0%とし、色 19 として表現した。明度と彩度の値は 20%,40%,60%,80%,100%を用いた。また、黒色のみ、明度値は 0%,20%,40%,60%,80%とした。

(6) 描画された絵の面積

絵の凸包の面積を特徴量としたものである。この特徴量は絵の大きさを表す特徴量であり、描かれたピクセル数を把握するためのものではない。

(7) 描かれたピクセル数が最も多い位置

キャンバスを 4×4 分割したときに塗られたピクセル数が最も多い位置を $x=(1,2,3,4)$ と $y=(1,2,3,4)$ で表す特徴量である。

(8) 分割されたキャンバスに描かれたピクセル数の割合

キャンバスを 4×4 分割したときに、各位置 $x=(1,2,3,4)$ と $y=(1,2,3,4)$ における塗られたピクセル数の割合を表す特徴量である。

(9) HOG(Histogram of Oriented Gradients)特徴量

HOG[11]とは輝度勾配のヒストグラムを用いて物体の形状の特徴ベクトルを表現するアルゴリズムであり、物体認識に使用さ

れる。HOG を用いて一筆描きされた絵の形状の特徴ベクトルを使用する。本研究では、HOG 特徴量に主成分分析を適用して特徴量を削減し、20次元としたものを用いた。

2.5 キャンバスの分割範囲

キャンバスの実際のサイズは 1000×1000[pixel]であるが、表示サイズは 800×800[pixel]である。特徴量(4), (7), (8)を求めるときにキャンバスを 4×4 分割するため、図 4 のように 800×800[pixel]の範囲を4等分割するような分割を行った。外側の正方形が 1000×1000[pixel]のキャンバスを、内側の塗られた正方形が表示している 800×800[pixel]のキャンバスを表す。x と y はキャンバスを分割したときにそれぞれの分割領域に割り当てられた位置を表す。

3. 結果

3.1 描画データ収集実験

機械学習で用いる大量の描画データを実験により収集した。実験では絵を描くための機材として液晶タブレット(Wacom 製 Cintiq13HD)を使用し、作成したイラストツールを用いて感情を表す絵を 1 色の選択と一筆描きによって描いてもらった。

絵を描いてもらう際、10 秒の制限時間を設けた。これは描き始めてからの時間であり、描く前に感情の絵を考える時間は制限していない。制限時間を設けるのは絵の複雑化を防ぐことと、人ごとの絵を描く時間差を小さくするためである。絵の描き直しは何度も行えるようにした。

収集実験では感情ごとに感情の強さを大・中・小と変えて描画データを取得することにより、一人から取得する描画データの数を増やした。一人から取得する描画データの数は 24 個となる。実験には 35 人が参加し、840 個の描画データを取得した。

参加者に対する感情の種類の説明において、受容と期待の 2つの感情については別の言葉と同時に提示した。この 2つの感情はどのような感情なのか把握しづらいと考えたため、受容は信頼と信用、期待は予期というように別に言葉に言い換えたものと同時に実験の参考資料として配布した紙に記載し、口頭での説明を行なった。

3.2 特徴選択

実験により取得した特徴量に対して特徴選択を行った。本研究では Weka(機械学習ソフト)を使用してカイ二乗独立性検定を用いた特徴選択を行った。

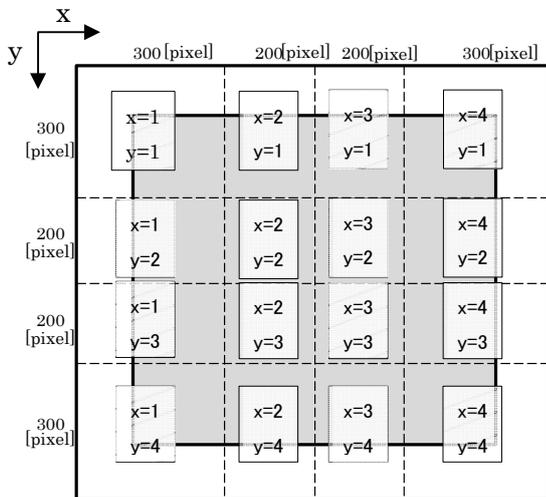


図 4. キャンバスの分割サイズ

本研究で使用した特徴量に対して特徴選択を行った結果、以下の特徴量が感情の認識に有効であることが分かった。

- 色(色相, 明度, 彩度)
- HOG 特徴量
- 描く方向の変化数
- 描き始め, 終わりの位置
- 筆圧の変化数
- 描かれたピクセル数が最も多い位置

3.3 SVMによる学習結果

(1) 8 分類における分類正解率

特徴選択によって選出した特徴量を用いて機械学習を行った。10 分割交差検定により評価したところ、分類正解率 51.9%を得ることができた。このとき出力された混同行列を表 2 に示す。この結果から、悲しみ, 受容, 怒りのそれぞれの感情については高い分類正解率を得ることができたことが分かる。しかし、喜びは期待に、嫌悪は恐れに、恐れは悲しみと嫌悪に、期待においては喜びに誤認識されやすいという結果となった。驚きに関しては、分類正解率が際立って低く、その他の感情に誤認識されやすいことが分かる。驚きは特に喜びに誤認識されやすい。

(2) 4 分類における分類正解率

本研究で使用した 8 分類の感情を 4 分類にまとめ、機械学習を行った。4分類の感情はポジティブ(喜び・受容), ネガティブ(悲しみ・嫌悪・恐れ・怒り), 驚き, 期待とした。驚きと期待に関しては肯定的, 否定的にも取ることができると考えたため、ポジティブとネガティブには含めないこととした。

感情を 4 分類としたときに各特徴量に対して特徴選択を行った結果、8 分類の感情を用いた場合での特徴選択の結果と比較したとき、認識に有効であると考えられる特徴量が変化した。特徴量の描く方向の変化数や筆圧の変化数は 4 分類の感情の認識においては有効ではないという結果が出たため、これらの特徴量は使用しないこととした。

10 分割交差検定により評価したところ、分類正解率 71.9%を得た。このとき出力された混同行列を表 3 に示す。この結果からポジティブとネガティブに関しては高い分類正解率を得たのに対し、期待はポジティブに誤認識されやすく、驚きは分類正解率がかなり低く、その他の感情に誤認識されやすいという結果となった。

3.4 人間の分類正解率

本実験で収集した一筆描きの絵を人に見せたときの分類正解率を求める実験を行った。この実験では、ランダムに全参加

表 2. 8 分類における混同行列

		出力感情								合計
		喜び	受容	悲しみ	嫌悪	恐れ	怒り	驚き	期待	
入力感情	喜び	46	10	0	0	1	7	11	30	105
	受容	15	65	3	5	3	3	4	7	105
	悲しみ	0	2	84	9	7	1	2	0	105
	嫌悪	1	2	11	48	35	5	3	0	105
	恐れ	1	1	20	24	49	4	5	1	105
	怒り	2	2	1	8	9	80	3	0	105
	驚き	30	10	4	8	10	16	16	11	105
	期待	32	11	5	0	0	1	8	48	105

表 3. 4 分類における混同行列

		出力感情				合計
		ポジティブ	ネガティブ	驚き	期待	
入力感情	ポジティブ	149	29	9	23	210
	ネガティブ	6	410	1	3	420
	驚き	48	40	7	10	105
	期待	51	9	7	38	105

表 4. 人間の認識における混同行列

		人間が絵から認識した感情								合計
		喜び	受容	悲しみ	嫌悪	恐れ	怒り	驚き	期待	
絵に含 ま れ る 感 情	喜び	23	3	0	1	0	2	3	4	36
	受容	5	14	1	3	3	0	2	8	36
	悲しみ	0	3	27	5	0	0	0	1	36
	嫌悪	0	2	9	12	7	0	4	2	36
	恐れ	2	0	8	9	12	0	2	3	36
	怒り	0	0	2	3	3	25	2	1	36
	驚き	0	2	1	4	1	4	17	7	36
	期待	6	7	4	5	0	1	5	8	36

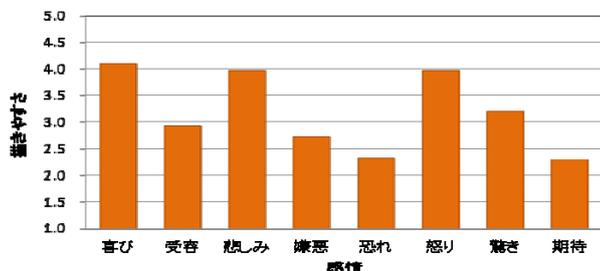


図 5. アンケート結果

者から 12 人を選び、その中から感情の強さごとに 4 人ずつの画像データを使用した。選出した計 96 個の画像データをランダムに並び替え、実験参加者に提示し、個々の画像データに対して認識した感情を記録した。計 3 人の実験参加者による実験の結果、分類正解率の平均は 47.9%となった。この結果から、人間に比べて、学習した認識器の分類正解率は高いものであることが分かる。しかし、実験参加者の数が少ないため、参加者数を増やしてデータを取得する必要がある。人間が一筆描きの絵から感情を認識したときの混同行列を表 4 に示す。

3.5 アンケート結果

描画データ収集実験終了後にアンケートを実施した。このアンケートでは、感情ごとに描く際の表現のしやすさを 1: 困難, 5: 容易とした 5 段階から選んでもらった。感情ごとの平均は図 5 のようになった。これから、描く際に表現しづらかった感情は恐れと期待であり、表現しやすかった感情は喜びと悲しみ、怒りであることが分かる。

3.6 考察

特徴選択の結果から、色はその他の特徴量に比べて感情の分類正解率に大きな影響を与えていることが分かった。また、感情の分類を 8 分類から 4 分類に変更することで分類正解率が向上したのは、混同されやすい感情を一まとめにすることができたためであると考えられる。

感情の 8 分類と 4 分類の共通の結果として、驚きの感情の分類正解率が極端に低い。人間による感情分類の結果を見ると、驚きは期待や恐れに比べて正解率が高い。描画データ収集実験の参加者に対するアンケートでも、驚きは期待や恐れなどと比較すると表現しやすいという結果が得られている。これらの結果から、驚きの分類正解率は向上の余地があると考えられる。驚きの分類正解率を向上させるために必要な新たな特徴量を発見することは今後の課題である。

4. おわりに

本研究では、描き手が一筆描きの絵で表現した感情を認識する感情認識器の開発を行った。感情の分類は Plutchik の分類(喜び, 悲しみ, 受容, 嫌悪, 怒り, 恐れ, 驚き, 期待)の 8 分類を使った。描画データから抽出した特徴量を使って、サポートベクターマシンにより描画データから感情の認識器を機械学習

した。描画データは、35 名から収集した 840 個の描画データを用いた。

特徴選択の結果から、感情認識には次の特徴量が有効であることが分かった。

- 色(色相, 明度, 彩度)
- HOG 特徴量
- 描く方向の変化数
- 描き始め, 終わりの位置
- 筆圧の変化数
- 描かれたピクセル数が最も多い位置

これらの特徴量を用いて 10 分割交差検定により評価したところ、8 分類の感情では分類正解率 51.9%を得ることができた。また、8 分類の感情をポジティブ(喜び, 受容), ネガティブ(悲しみ, 嫌悪, 怒り, 恐れ), 驚き, 期待の 4 分類にまとめ、10 分割交差検定により評価したところ、分類正解率 71.9%を得ることができた。感情の分類を 4 分類とすることで分類正解率が向上したのは、混同されやすい感情を一まとめにすることができたためであると考えられる。

参考文献

- [1] Justine Cassell, Joseph Sullivan, Scott Prevost and Elizabeth F. Churchill eds., "Embodied Conversational Agents", MIT Press, 2000.
- [2] 見館好隆, 館野泰一, 脇本健弘, 望月俊男, 宮田祐子, 中原淳, 三宅なほみ, "ロボットによる主体的な発話支援の有効性について", 日本教育工学学会誌, Vol.37, No.3, pp.209-227, 2013.
- [3] Kohji Dohsaka, Ryota Asai, Ryuichiro Higashinaka, Yasuhiro Minami, and Eisaku Maeda, "Effects of Conversational Agents on Human Communication Agents on Human Communication in Thought-Evoking Multi-Party Dialogues", IEICE Transactions on Information and Systems, Vol.E97-D, No.8, pp.2147-2156, 2014.
- [4] 山下真知子, "感情からイメージされる色彩について: 高齢者の回復期ケアを目的とした施設空間の色彩設計に関する研究", 大手前大学論集, pp.289-316, 2008.
- [5] Michela Dellagiacoma, Pamela Zontone, Giulia Boatom, and Liliana Albertazzi, "Emotion Based Classification of Natural Images", in: Proceedings of the International Workshop on DETecting and Exploiting Cultural diversity on the Social Web (DETECT'11), pp.17-22, 2011.
- [6] Paul Ekman, Wallace V. Friesen 著, 工藤力訳, "表情分析入門—表情に隠された意味をさぐる", 誠信書房, 1987.
- [7] Robert Plutchik, "The nature of emotions", American Scientist, Vol. 89, Iss. 4, pp.344-350, 2001.
- [8] 福井康之, "感情の心理学: 自分とひととの関係性を知る手がかり", 川島書店, 1990.
- [9] Nello Cristianini., John Shawe-Taylor 著, 大北剛訳, "サポートベクターマシン入門", 共立出版, 2005.
- [10] 阿部重夫, "パターン認識のためのサポートベクトルマシン入門", 森北出版, 2011.
- [11] Navneet Dalal, Bill Triggs, "Histograms of oriented gradients for human detection", Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.886-893, 2005.